

ED 022 099

By-Guilford, J.P.

INTELLECTUAL DECLINE.

Pub Date 67

Note-26p.; Chapter 18 in THE NATURE OF HUMAN INTELLIGENCE by J. P. Guilford, published by McGraw-Hill Book Co., New York, N.Y.

EDRS Price MF-\$0.25 HC-\$1.12

Descriptors-*AGE, CONTINUOUS LEARNING, HEREDITY, *INTELLIGENCE, INTELLIGENCE TESTS, MEMORY, *MENTAL DEVELOPMENT, *OLDER ADULTS, PHYSIOLOGY, RESEARCH, THOUGHT PROCESSES

In investigations of decline of intellectual status with age, cross-sectional and longitudinal studies give divergent results--the former show almost universal declines in test performances among older groups while the latter often show gains. Although few structure-of-intellect (SI) factors have been explored in relation to adult ages, two kinds of SI abilities stand out with respect to defects in the elderly--the formation and use of systems of information and the ability to be flexible in shifting between classes and in transformation of information. The older person's capacity for information intake and for handling information in productive thinking may be limited, perhaps due to a weakness in short-term memory. A rather general weakness is susceptibility to interference and inability to correct errors. Little is known about the causes of intellectual decline. Heredity appears to have a bearing, possibly largely because it affects development. Inadequate oxygen supply to the brain is a suspected cause of decline. Exercise of brain cells through continued education, formal and informal, including that provided by occupational activity, may have halting effects, but this hypothesis needs better experimental investigation. (aj)

The Nature of Human Intelligence

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

J. P. Guilford
Professor of Psychology
University of Southern California

"PERMISSION TO REPRODUCE THIS
COPYRIGHTED MATERIAL HAS BEEN GRANTED

BY Ray H. Poe
McGraw-Hill

TO ERIC AND ORGANIZATIONS OPERATING
UNDER AGREEMENTS WITH THE U.S. OFFICE OF
EDUCATION. FURTHER REPRODUCTION OUTSIDE
THE ERIC SYSTEM REQUIRES PERMISSION OF
THE COPYRIGHT OWNER."

McGraw-Hill Book Company

New York St. Louis San Francisco Toronto London Sydney

AC 002 496

ED022099

The nature of human intelligence

Copyright © 1967 by McGraw-Hill, Inc. All Rights Reserved.
Printed in the United States of America. No part of this publication
may be reproduced, stored in a retrieval system, or transmitted, in
any form or by any means, electronic, mechanical, photocopying,
recording, or otherwise, without the prior written permission of
McGraw-Hill, Inc.

Library of Congress Catalog Card Number: 67-11207

25135

1 2 3 4 5 6 7 8 9 0 M P 7 4 3 2 1 0 6 9 8 7

18

Intellectual decline

As in the preceding chapter, we are concerned with intellectual abilities in relation to age, but here we are concentrating on the ages at which typical individuals reach their primes and the ages beyond which they lose in measured abilities. The rates of decline with age and for different abilities have both social and theoretical interest. We shall look into some of the conditions that are known to have some bearing upon rate of decline and shall be on the lookout for general principles of decline. Before we consider the known facts and theories about decline, however, it will be necessary to pay some attention to methods of investigation, for the results cannot be properly interpreted and evaluated without considering how they came about.

Research methods

As in the investigation of growth of abilities, there are two major strategies, cross-sectional and longitudinal. Neither approach is fully satisfactory, and both have weaknesses that need to be kept in mind. As before, some of the difficulties pertain to sampling problems and changing environmental conditions, and others pertain to tests and testing conditions. Since very much of the research has involved scores from separate tests, whether they have been designed as measures of intellectual-aptitude factors or not, it is important to consider the factorial nature of tests used with the aged and middle-aged groups. Factor structures for elderly populations is of some theoretical interest.

Sampling problems Many of the same sampling problems are encountered whether the cross-sectional or the longitudinal approach is taken. Certain conditions that are related to measured abilities change systematically with age. Conditions such as education, its quantity, its quality, its kind, and the time since formal education of any consequence has ceased, income and what opportunities money can purchase, and occupation constitute the major potential environmental determiners of what individuals can do on tests. They contribute to individuals not only different supplies of information for memory storage but also exercise in skills for handling that information, skills that are also represented in memory storage preparatory to test performance. The time elapsing since rehearsal is also unusually relevant and generally is not controlled. Ideally, one should want stratified-random sampling, as in a Gallup poll, with control of all pertinent variables. Only one such study has been reported (R. L. Thorndike & G. H. Gallup, 1944).

There is also the problem of selective survival. It has been found with the use of twin subjects past the age of sixty that the twins who died during a nine-year interval following the first testing had averaged significantly lower than their twin mates in five Wechsler tests, Digit Symbol, Substitution, Block Design, Similarities, and Vocabulary (Jarvik, Feingold, Kallman, & Falek, 1962). They were not lower on other tests. If the same kinds of differences apply to groups tested in younger years, the variable of longevity is potentially even more widely relevant. Surviving populations are apparently not like nonsurviving populations in every intellectual ability. It is obvious that we can never know what some individuals who did not survive might have done on tests had they lived. Nor

can we select for testing at earlier ages only those who are destined to survive to the age of eighty or whatever year applies to the oldest group used in a study.

Of the various possible confounding variables, Schaie (1959) found that the most relevant ones were family income, education, and occupation, which were most closely related to test scores after the age of fifty. But when Schaie obtained a multiple-regression equation predicting test scores from those three demographic variables and adjusted the scores, taking into account these predictions, he found that the age differences in PMA scores remained much the same as before adjustment.

Although the correlations between these variables and test scores may be negligible, as in Schaie's study, it is well to know whether they are high enough to be of some consequence in particular studies. As an example of such correlations, Bilash and Zubek (1960) found coefficients ranging from .20 to .42 between amount of education and test scores, with age partialled out. For a large sample ranging from twenty-five to sixty-four years of age, Birren and Morrison (1961) found that the correlations between the 11 WAIS tests and age ranged from $-.02$ to $-.46$, with a mean of $-.21$. With education held constant, the range was $+.22$ to $-.38$, with a mean of $-.07$. A greater effect of education might well have been apparent if the age range had been extended to eighty. On the whole, because of the positive correlations between education and test scores (Birren and Morrison report WAIS correlations of .40 to .66 with a mean of .51, with age not partialled out) and the negative correlation between education and age ($-.29$, from the same source), the effect would have been to show slightly more decline in scores as functions of age than would have occurred with education controlled.

Besides the possible application of analysis-of-covariance procedures, as Schaie (1959) suggests, if samples are sufficiently large, groups could be fractionated according to educational level, as Pacaud (1955) did with 4,000 French railway employees. With two levels for education, she did find the same rates of decline in test scores in the two groups. But she had the assurance that educational level had been controlled. With other kinds of tests that are more strongly related to education, the result might have been different. Later in the chapter some instances of effects of occupation will be mentioned.

In the study of growth, institutional life has appeared to be of substantial importance in some instances, depending upon the kind of environment provided. The same may be true of studies of the elderly. If stimulation of the brain is a significant contributor to development, it may also be important for maintaining intellectual status already achieved. Residence for a number of years in Veterans Administration hospitals and in rest homes for aged should be taken into account. If continued stimulation of the brain is needed to keep the intellectual level from declining, the elderly who "just vegetate" must surely lose intellectually.

At the higher ages, particularly past sixty, one also has to consider the possibility of degenerative diseases, such as arteriosclerosis, that might be contributing to decline. Usually investigators have been careful to distinguish between normally aging subjects and others with organic degeneration. Even the condition of general physical health should be considered, for Birren, Botwinick, Weiss, and Morrison (1963) have demonstrated some significant differences in mean scores of elderly subjects with good versus poor physical health.

Tests and test conditions Both Wechsler (1958) and H. E. Jones (1959) have raised some pertinent questions concerning appropriateness of tests for older individuals. Since testing as a social institution has been increasing in its application in the general population, younger persons are more experienced and more familiar with tests of different kinds

than the elderly. There is no question of the general benefits from having had such experiences. The variable of test-wiseness should contribute to more apparent decline in cross-sectional studies.

In longitudinal studies, the effects of taking the same test or a different form of the same test on repeated occasions are contaminating biases. As we shall see later, longitudinal studies, with retesting after one to thirty years, are likely to yield increases in means of scores, whereas comparing age groups in cross-sectional studies is likely to show decreases with aging in almost all tests. How much of this can be attributed to retesting is hard to say. Some of this kind of result may be incidental to the fact that in longitudinal studies investigators have happened to use tests heavily weighted with verbal content, emphasizing factor CMU, in which increases are common even in cross-sectional studies. But the aged show increases in scores in other kinds of tests also, as Kamin (1957) has found. Two groups with mean ages of about seventy-two were tested four times with Thurstone's PMA test. The gains in scores were attributed not to specific memory of items but to more general skills in test taking. Such effects of practice tend to wear off with time, as in normal forgetting, and if there are several years between testings, such effects should not be so important in comparing results related to age.

Another question often raised pertains to motivation. Many of the tests were designed to appeal to children and young adults. How well will they also appeal to older adults and to the elderly, particularly if some of the latter happen to be teachers and university faculty members, as in two studies (Garfield & Blek, 1952; Sward, 1945)? Mature individuals are differentiated with regard to occupations, also, differing from school populations in this respect. Tests that appeal to a mechanic may not appeal to an artist or to a businessman.

There is also the problem of motivation for competing in tests, regardless of their contents. Defensive middle-aged and elderly individuals will not volunteer and will not compete. Those who are not sure of themselves and who have pride will not enter into studies voluntarily. Subjects who enter without volunteering will come to the testing with different degrees of cooperation. This condition is better controlled in individual testing, but group testing is the expedient way and therefore tempting to use. Older subjects, less sure of themselves, are less likely to guess, which may be a handicap in multiple-choice tests. Kamin (1957) found that a little extra inducement, in the form of offering a prize for the most improvement shown in taking the PMA tests a fourth time, was followed by marked increases in scores in some of his aged subjects. Some Ss who had not gained in the second and third administrations gained suddenly. The most improvement was apparent in the Word Fluency test, a highly speeded test. Even under normal conditions, such a test may reflect some variance in a motivational variable.

Speed as a test condition It is well recognized that with advancing age there is a general slowing down in almost all respects. This tendency is obvious in the natural ways of doing things, and it shows up in tests that demand rapid work in order to achieve a good score. Particularly since Lorge (1936) emphasized the problem, investigators of age in relation to abilities have given it considerable attention.

Lorge demonstrated the principle with minimal evidence. He administered to subjects ranging in age from twenty to seventy three intelligence scales that varied with respect to degree of speeding: a form of the CAVD,¹ which allows unlimited time; the Army Alpha, which has moderate time limits for its parts; and the Otis twenty-minute test,

¹ A test title using initials for its components: Completions, Arithmetic, Vocabulary, and Directions.

which is highly speeded. The correlations of scores with age were $-.27$, $-.36$, and $-.48$ for the three tests, respectively. Lorge concluded that decline curves are somewhat exaggerated in the losses that they show for intelligence, in spite of the fact that the CAVD test correlated negatively with age to about the same degree as the Army Alpha. Tests of the Army Alpha type with respect to speeding have usually been employed in deriving decline curves. Lorge went much too far in concluding on the basis of his results that intellectual *power* probably does not decline with age, particularly when the CAVD test did show some negative correlation.

In order to gain a better idea of how much speed conditions for some tests bias the decline curves, some investigators have given the same tests, for example, the Wechsler tests, which are widely used in studies of aging, under normal time limits and with unlimited time. Dibner and Cummins (1961) gave the Wechsler tests in the two ways except for Digit Symbol, which would probably yield all perfect scores when given without a time limit. They found little gain for the unspeeded condition, with a mean gain of 2 points in IQ for the performance battery and a mean gain of 1 point for the total IQ.

Doppelt and Wallace (1955) gave five of the Wechsler tests to subjects of sixty and older with and without time limits. The "power" scores averaged 5 percent higher, but the means were never so much as 1 unit higher. With the PMA tests, Schaie et al. (1953) found little difference in the decline curves whether the regular or the untimed scores were used. It does not appear, then, that in two of the popular scales used in studies of normal aging the speed feature is much of a handicap to the elderly or biases the decline curves materially. This does not justify our generalizing the same conclusion to all kinds of tests, of which there are a great many not represented in these two scales, or generalizing to pathological populations.

There are some tasks in which speed has been shown to matter. For example, Chown and Heron (1965) cite an instance in which elderly subjects make better learning scores when they pace themselves in memorizing. This is not true of younger subjects. Systematic addition to exposure time for the elderly also gives them some advantage in learning. The elderly seem to need much redundancy of input, which the extra time allows them.

Botwinick, Brinley, and Robbin (1959) proposed a "modulation" hypothesis to the effect that the elderly are not so adaptable to speeds other than those that they prefer or that come naturally to them. Elderly subjects (mean age of seventy-two) could neither increase nor decrease their writing speeds as much as young subjects (mean age of twenty-two) could under instruction to do so. There are some who might cite these results as another indication of lowered flexibility, but the kind of flexibility would need to be specified; there are several kinds.

In the study just cited, the normal writing speed was slower for the elderly, the mean writing time for a certain task being fifteen seconds, as compared with eleven seconds for the young group. Birren (1955) questioned whether differences in writing speed might be a factor in a test such as addition. He obtained correlations between writing speed and scores in an addition test. The correlation between writing-speed scores and addition scores was highest if the addition problems each contained 2 digits, when relatively more of the working time was spent in writing, and dropped systematically as the number of digits per problem increased to 10, when the least writing was involved. Since this decrease in correlation was about the same for the elderly as for the young, Birren concluded that writing speed did not "... disproportionately impair the output of the elderly" in addition tests. He suggested the hypothesis that the slower adding speed of the elderly reflects a slowing of their perceptual processes.

There are other studies that bear upon speed and aging. The few examples given

show that the speed variable needs to be investigated in connection with each kind of task, and the source of the slowing of the elderly needs to be pinpointed if we are to derive some really useful general principles on the role of speed.

Speed as a test control On the general role of speed in aptitude tests, there are some theoretical points that should be mentioned. There are certain tests, like Digit Symbol, that would be pointless without time limits such that almost no one can complete all items, for errors could be zero and all scores maximal if liberal time were allowed. There are some abilities for which speed is an essential condition, for example, tests of fluency, in which speed of recall, or retrieval of information, is an essential aspect of the aptitude.

In other instances, time limitation is a condition that is needed in order to control the way in which the examinees work on the items. If given too much time, they can devise strategies that involve operations that change the nature of the test. It would then measure an ability or abilities for which it was not intended. Under the speed condition, the person who is high in the ability that the test is designed to measure, visualization, for example, can do the items successfully, thus emphasizing his visualizing function. Under liberal time, the person who is low in visualizing ability might reason out the answers and make a good score, which would be misleading as to what is measured. In other words, speed is often an important experimental control where other controls are less effective. In a vocabulary test, liberal time can be given because there is only one way a good score can be obtained: the examinee knows the meanings of the words. Determining the optimal timing for tests for the elderly, optimal in terms of measuring the intended ability, presents a technical problem of which there has been no apparent investigation.

Factor structure at different age levels There are two important reasons for wanting to know about the factor structure of tests at different age levels among adults. One is the question of whether the aptitude factors, once developed in differentiated form, remain so throughout the ordinary life-span. There has been no prominent general hypothesis about this question as a counterpart to the Garrett hypothesis that applies at the lower end of the age scale. The other reason is a question of methodology. Investigators of decline in different abilities sometimes merely question whether a certain kind of test measures the same factor or factors at all age levels. It is important to know whether this is so if we are not to be led astray by decline curves, which might represent one factor through the age of fifty and another factor thereafter. One should not expect such sharp transitions as this statement implies, but there could be relative shifts in factor content as a function of age.

Evidence from factor analysis There have been attempts to answer the question insofar as the 11 Wechsler tests are concerned. As the writer has stated earlier, the tests of the Wechsler scales are poor material for factor analysis if they are analyzed alone, for probably as many SI factors are represented as there are tests, perhaps more. The outcome has usually been a verbal factor, a performance factor, a weak memory factor, and sometimes as many as three additional factors. There is a general, or *g*, factor where that is demanded. The verbal and performance factors are, of course, composites, roughly representing the difference between certain combinations of semantic and figural abilities.

Analysis of the 11 Wechsler tests by themselves can do one thing: it can tell us whether there is general stability of intercorrelations at different age levels, and if there is, we may conclude that the verbal and performance scores represent rather stable composites. We can draw no conclusions regarding any more refined stabilities of factor assess-

ment that the tests measure singly. One qualification should be stated, lest the impression be left that the two standard composites, verbal IQ and performance IQ, can be accepted as purely semantic and figural. There is a semantic-ability test within the performance group, Picture Arrangement, which should measure factor NMS, the convergent production of semantic systems. And there are symbolic tests, Memory Span and Digit Symbol, in both composites. These two are, indeed, memory tests, but they represent two symbolic-memory abilities. They might be scored as a third composite, leaving the other two less complex.

Where factors for the Wechsler tests are found at all testable age levels, interpretations of factors are much the same (Birren, 1952; Cohen, 1957; Green & Berkowitz, 1964; Riegel & Riegel, 1962). One finding that may be of some significance is the fact that fewer interpretable factors sometimes appear at the oldest levels. This trend was clearest in the study by Green and Berkowitz (1964). The trend suggests some simplification of factor structure for the oldest tested groups, usually above seventy, in a possible retreat from earlier differentiations.

Evidence from intercorrelations Some of the intercorrelation data support this trend, but one study does not. Green and Berkowitz (1964) found that the averages of correlation coefficients increased with age, from .41 in the youngest group (under twenty-nine) to .51 in the oldest (over sixty-five). Kamin (1957), using the PMA tests with samples of only 25 each, found the following ranges of coefficients in three different groups:

<i>Kind of group</i>	<i>Range of coefficients</i>
Normal aged (mean age seventy-two)	.20-.56
Institutionalized aged (mean age seventy-two)	.13-.59
High-school level	-.35-.49

The discordant results were reported by Schaie et al. (1953), who used the PMA tests with subjects aged fifty-three to seventy-eight. The correlations ranged from .06 to .31, which were lower than those cited for young adults by Thurstone and others. The reliabilities were as high as are usually reported. The unusual range of ages should lead one to expect substantial correlations because of some declines with age common in most of these tests. The ranges of correlations reported by Kamin and Schaie et al. are promising for the possibility of separation of factors in the normal aged.

Intercorrelations for senile psychotics When the population is diagnosed as senile, the intercorrelation picture is quite different. From such information and other sources, Dörken (1954) concluded that senile dementia is not merely a more rapid pace of aging such as is found in the normal aged. For the senile population he reported an average correlation of .63. Some data to which the writer has had access provide similar results.¹ With a sample of 360, five-sixths of whom were suffering from various degrees of senile dementia, the intercorrelations of 11 tests ranged from .44 to .81, most of them being above .60. When the total group had been divided into a higher and a lower subgroup on the basis of a rating for intellectual status derived independently of the tests but correlating highly with them, in each group the intercorrelations were factor-analyzed. Because of the generally high correlations, axes were rotated obliquely. For the higher group, the

¹ The writer is indebted to Oscar J. Kaplan for the use of these data.

average of the cosines between pairs of the six primary vectors was .58. For the lower group, the corresponding mean was .67. Seven factors were extracted in the latter case, two experimental variables having been added to the analyzed battery.

It is fairly clear that the factor structure is rather different for senile and normal adult populations. To cite a particular instance, the correlation between a vocabulary test and a number-skills test was .78 for the combined sample of 360. There would be no chance of separating a verbal from a number factor, as would be true in young adults and possibly in normal aged adults. Hallenbeck (1963) concurs in the belief that with brain damage the elderly do differently on tests than the normal elderly, for which he gives evidence. The general increase in intercorrelations can be attributed to organic deterioration that affects abilities in the same direction. There are other, less systematic changes, apparently, for Dörken (1954) concluded that for seniles tests correlate zero with age, whereas for normals they ordinarily correlate negatively. Only by analyzing the same battery of tests for SI factors in the senile, the normal elderly, and normal adult samples can we make good comparisons of factor structures. It may be that the senile population would show little more than a general factor, with few other systematic underlying variables of ability, depending upon the degree of deterioration.

Whatever increasing intercorrelation is found in the *normal* aged can be attributed to other possible causes, as well as to organic conditions that may affect different abilities similarly. For example, Welford (1958) has pointed out that the normal elderly have a remarkable capacity for organizing their behavior strategically so as to compensate for declining abilities. Such a capacity might affect scores in many different tests in similar ways and thus contribute to increments in correlations.

Decline in composite scores

Although decline curves from different tests are more strikingly different than are growth curves, thus calling for a more analytical approach to the study of decline, there has been some interest in decline as indicated by single scores from IQ tests. The Wechsler scale for adults has been the favored instrument for this kind of study. Somewhat different conclusions are reached from cross-sectional versus longitudinal approaches to the problem.

Decline shown by cross-sectional methods Wechsler (1958) presents a decline curve for total score from his scales, based upon very large samples, as reproduced in Figure 18.1. The scale for ability is in terms of standard scores with the mean of zero at the highest point, which comes near the age of twenty-five. The decline after the age of thirty is almost linear to the age of seventy-five. It should be remembered that this curve represents performance in a particular composite of tests, on a certain kind of scale. Since the scale does not include all intellectual factors, it should not be regarded as representative of all intelligence.

Wechsler also provides decline curves for his verbal and performance scales separately. The former shows less rapid decline than the total scale, and the latter shows more rapid decline. These two curves are not shown here, because there is much more to be learned from information concerning declines in particular tests. Wechsler tests showing the most rapid declines are Digit Symbol, Picture Arrangement, and Block Design. The tests showing the least rapid declines are Information, Vocabulary, and Comprehension, all semantic tests, the first two dominated by factor CMU and the third by CMS and CMU.

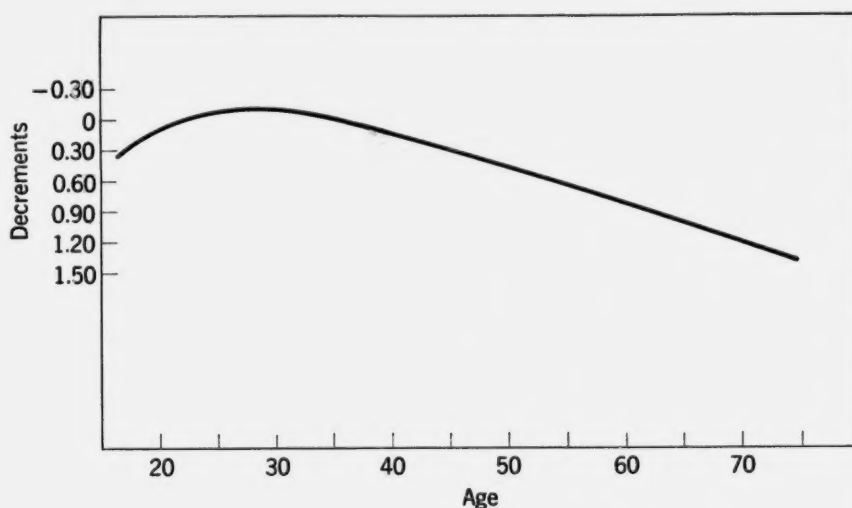


Fig. 18.1 Curve representing intellectual decline in terms of IQ-test scores. (From Wechsler, 1958. Reproduced by permission.)

The Wechsler deterioration quotient Wechsler has taken advantage of the differential rates of decline of his separate tests to derive from combinations of them a score known as a deterioration quotient (DQ). This is a ratio derived from the "hold" tests, those showing little decline, and the "don't-hold" tests, those showing much decline. The hold combination includes Vocabulary, Information, Object Assembly, and Picture Completion. The don't-hold combination includes Digit Span, Similarities, Digit Symbol, and Block Design. The DQ is given by the formula

$$DQ = \frac{\text{hold} - \text{don't hold}}{\text{hold}}$$

where standardized composite scores are used.

The basic assumption is that normal aging is a manifestation of deterioration. The application of the DQ index, however, is designed to indicate deterioration attributable to pathological origins. This use involves a very serious risk. In an earlier discussion, evidence was cited for the fact that the effects of normal aging upon tests are not the same as the effects of senile dementia and that deterioration in dementia produces different and perhaps unpredictable effects in test performances. For example, in some individuals, verbal ability might be one that declines drastically (Inglis, 1958).

The greatest source of error of interpretation of the index, however, is that a ratio may simply indicate how strong the individual is in one set of factors relative to another set. The mean of the DQ values from a population is approximately zero at all ages above sixteen, and standard deviations are just about the same at all ages (Wechsler, 1958). This would mean that the DQ index should show many young deteriorated cases. Negative DQ values would also be difficult to interpret. What is the opposite of deterioration? The older person who shows strong deterioration according to Wechsler's DQ index might have had such a DQ score most of his life. This would be true if his profile of WAIS scores remained fairly constant over the years.

In terms of discrimination of senile from nonsenile aged subjects, the Wechsler DQ has done poorly. For example, Botwinick and Birren (1951) found that 31 senile patients were not significantly different from a control group of 50. Total IQ on the WAIS, however, discriminated the two groups significantly.

Longitudinal studies Longitudinal studies of changes in general intellectual status of adults more often than not actually show gains in scores. The study with the longest delay interval was done by W. A. Owens, Jr. (1953). Owens retested in 1950 with the Army Alpha Examination 127 men who had had the same examination as college students in 1919, with an interval of thirty-one years. The average change in score was a gain equal to $\frac{1}{2}$ standard deviation. This gain could be attributed entirely to gains in four verbal tests, in which the gains varied from 0.5 to 0.9 standard deviation. There was no significant loss in any test. Younger subjects tended to make greater gains, as did those who had attended college more than five years as compared with those who had attended less than four.

In evaluating these findings, there should be some reservation because of a sampling feature. The subjects retested were volunteers, and it may be questioned whether those who would not cooperate were among those gaining less or possibly losing. Because of the single long interval and because there was no testing at the time of graduation, we do not know when the observed gains occurred, whether during the years in college or in the interval after leaving college. We also do not know how much to attribute the gains to retesting, even after so long an interval.

Of the 127 in Owens's sample, 96 were retested again in 1961, eleven years after the first retesting (Schoenfeldt, Gillmer, Kelley, & Owens, 1963). By comparing means for three factors (verbal, number, and reasoning), it was found that the verbal mean had dropped slightly, the reasoning mean had held fairly constant, and the numerical mean, which had dropped slightly in the first retesting, had dropped more in the later interval.

Jarvik et al. (1962) tested many pairs of twins who were past sixty, then retested them after one year and also after nine years, with some of the Wechsler tests. There were increases in means of scores for the first retesting, which the authors attributed to improvement with experience in being tested. There were decreases between second and third testings in all tests except Digits Forward. It would be desirable to know whether the same subjects would have gained as much if they had been retested immediately after their first testing. Owing to a probable increasing incidence of poor physical health during the seventh decade, it would be important to know whether the average status in this respect was the same in the third testing as in the second.

An exception to the general rule appeared in a study by Berkowitz and Green (1963). Their subjects were inmates of a Veterans Administration hospital, with a mean age of 56.3 at the time of first testing and 65.0 at the time of second testing. Decrements were found in all the Wechsler tests after the interval, which averaged about nine years. The slight but significant declines in IQ were about the same for both verbal and performance scales. It is possible that the institutional life of these individuals contributed to the losses.

Decline in intellectual factors

An example of how decline curves for particular tests differ is shown in Figure 18.2 for four of the Wechsler tests. Such differences are found even when the tests are not univocal for intellectual factors. The Information test is essentially univocal for factor CMU (verbal comprehension), Block Design is for factor CFT (spatial visualization), and Digit Symbol appears to be for factor MSI (memory for symbolic implications) (P. C. Davis, 1956). The Arithmetic test, like arithmetical-reasoning tests in general, may have two leading factor components, CMS (general reasoning) and numerical facility. Its decline curve then represents two or three factors in combination, which means that its rate of decline may be a compromise or an average, one component declining at a more

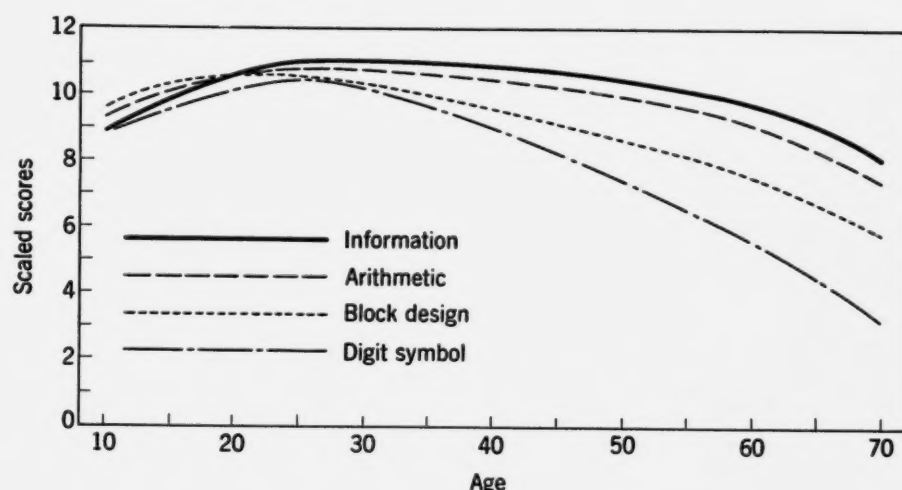


Fig. 18.2 Curves representing differential rates of decline in four Wechsler tests. (From Wechsler, 1958. Reproduced by permission.)

rapid rate than another. Again, we see the desirability of using univocal tests if we want unambiguous pictures of declining functions.

Tests of cognition Enough special tests have been used in studies of decline to make it worth our while to attempt to see whether there are any systematic relations of decline to structure-of-intellect categories. The evidence is very spotty, at best. At least the categories will serve as a basis for organizing the discussion.

Welford (1958) has made the comment that the apparent weaknesses of the elderly in learning and retention may actually be due to weaknesses in cognition; new experiences are not sufficiently impressed upon the individual's nervous equipment because input has not been well developed. The fact that the elderly often show very good retention and recall for earlier experiences also indicates that insofar as retention itself is concerned, there is not general impairment. But since there are quite a number of cognitive abilities, it is necessary to consider whether all or just certain ones are impaired.

Vocabulary Of all tests, those for CMU have shown the least decline with age. In a cross-sectional study, Tremblay (1964b) has reported that with a 150-item multiple-choice vocabulary test, the performance curve continues to rise at a decelerated rate to the age of sixty. R. L. Thorndike and G. H. Gallup (1941) found no change from age twenty to age fifty, with an average loss of only 1 point (in a 20-item test) to the age of sixty-five, a loss equivalent to ten months of mental age. Garfield and Blek (1952) gave a certain vocabulary test to teachers in training (ages twenty to thirty), teachers, and retired teachers, one group with ages forty to fifty and the other sixty to seventy. The means of scores were approximately 31, 35, and 38, respectively, all differences being significant. Even in cross-sectional assessments, such as all these examples have been, vocabulary continues to grow slightly, particularly in populations in which words are an important commodity.

The conclusion is sometimes drawn that "verbal ability" remains high, but generalizations beyond one semantic ability to others is not justified. Within the category of semantic-cognition abilities, in which there are six, there is evidence for decline in only one or two other than CMU. The decline in CMS (general reasoning), which is the leading

component of arithmetic-reasoning tests, is suggested by the curve in Figure 18.2. K. F. Riegel (1959) has demonstrated that a young group (mean age about nineteen) had a higher mean than an older group (above sixty-five) in a test that probably measures factor CMC. A sample item reads:

A GRANARY belongs with: field, stable, farm, barn, plough.

A recognition of classes must be a feature of such a test.

Some figural-cognition abilities There are a few indications of declines in figural-cognition abilities. Welford (1964) states that in tests involving the viewing of objects under difficult conditions the elderly make more errors. This description fits tests of factor CFU-V, the cognition of visual-figural units. For factor CFS-V, Schaie (1958) has reported that in Thurstone's PMA Space test there is a linear decline after the age of thirty-three. Bilash and Zubek (1960) found that the curve for a test of Space Relations kept to a high level to the age of forty-five, after which there was a rapid decline. The test may have had some involvement with factor CFT (visualization). Wechsler (1958) reports that the test Block Design, another CFT test, is one of three in his battery with most rapid decline.

Some symbolic-cognition abilities Information about decline in symbolic-cognition abilities is limited to two tests: King's Reasoning test, which has some items that should measure CSC and others that should measure CSS; and Thurstone's PMA Reasoning, which should measure CSS. The King test showed an almost continuous decline from twenty to seventy, with some arrest between twenty and forty-five (Bilash & Zubek, 1960). The Thurstone test showed a linear decline from twenty to sixty (Schaie, 1958).

Memory abilities A survey of the scattered experimental and psychometric information regarding memory of different kinds reveals conflicting pictures. There are no reports of actual evidence to the effect that any memory abilities improve with age, beyond the youthful period, but in some respects there is decline and in some respects not. Most of the studies need to be replicated and extended to other areas of information, and further studies in depth are needed for particular varieties of memory.

Short-term memory Discussions of what is called "short-term" memory in Chapter 13 adopted the view that it deals with a phenomenon lasting only a second or two after stimulation. The somewhat vivid persisting aftereffects of stimulation permit the individual to extend his capacity for input, his momentary capacity being very limited. By virtue of this overlapping of current experiences he is able to take in and perhaps to fixate more information than would otherwise be possible.

Experiments of the Broadbent type, in which digits are spoken into earphones in the two ears, three in one and three in the other, simultaneously or alternately in the two ears, have been applied to the elderly. It is found that the elderly do as well as younger subjects in reporting the set of three digits to be reported first but show relative defects in reporting the second set of three exposed digits (Welford, 1958; 1964). Although debate has continued over whether in younger subjects the loss of information is due to time delay or to interferences, with the elderly it looks as if interference were the key determiner. In previous chapters the view has been that interference can best be interpreted as confusion of information, which means loss of discrimination or loss of information. Thus, the increased intake capacity offered to younger subjects by virtue of short-term memory

seems to be more or less withheld from the elderly. Any such possible gain in capacity is offset by interferences, very likely more than offset.

Memory for units and systems There happens to be a little evidence regarding the memory for units and systems of information, figural, symbolic, and semantic. Memory-span tests, for series to be recited in the forward direction, should be at least in part measures of factor MSS. Gilbert (1941) reports only a moderate amount of decline. Cameron (1943) found that even seniles can remember three digits for several minutes but that interpolated spelling activity ruins recall for the digits. This episode of retro-active inhibition is obviously an interference effect, having its influence apparently beyond the short-term-memory interval. It is an interesting question to what extent the common use of telephone numbers, social security numbers, and the like in the United States culture has helped to keep memory-span ability from declining more than it otherwise would.

For a test of Design Memory, which should be a measure of factor MFU, Tremblay (1964b) has reported a curve that reaches its maximum at the age of seventeen and starts its decline after about thirty. In the study of memory by H. E. Jones, H. S. Conrad, and A. Horn (1928), in which subjects answered questions about facts presented in a motion picture, a performance curve in relation to age was very similar to those for composite scores of intelligence, as in Figure 18.1. The kind of memory involved should have been largely for semantic units, that is, for factor MMU.

Memory for implications The only memory-for-implications ability about which much is known is MSI. From at least two studies (P. C. Davis, 1956; de Mille, 1962) there is evidence that numerical-operations tests and the Wechsler Digit Symbol test both measure this ability to some extent. The former emphasize well-practiced implications of long standing in the case of adults, whereas the latter deals with very recently formed associations, so recent they can even be put in the category of short-term memory, as the examinee looks back and forth from code to items on which he is working.

With a numerical-operations test, Bilash and Zubek (1960) found that performance remained at about the same high level to the age of forty-five, then declined rapidly. Schaie (1958) reported from use of the PMA Number test a slight rise to about fifty, then a rapid decline. Wechsler (1958) reported the Digit Symbol test among the three with most rapid decline. His curve shows the decline starting somewhere between twenty-five and thirty years of age.

Birren, Allen, and Landau (1954) made a rather intensive study of performances on addition tests in ages from sixteen to ninety. There would be much risk in attributing their conclusions to factor MSI. Recent experience shows that an addition test has a relatively large specific component besides MSI variance (Tenopyr, 1966). Other findings (Petersen et al., 1963) show that numerical-operations tests are also loaded with NSI. Some of the features shown in Birren's experimental results could thus be possibly attributed to factor NSI, a second common-factor component of numerical-operations tests, or to the specific component in an addition test.

Birren et al. were concerned with speed and accuracy as functions of age and of the numbers of additions per item, also with interactions between age and the other condition. They varied the items with respect to numbers of additions from 1 to 25 (2 to 26 digits). The measurements taken were average time per item and proportion of correct item sums. One thing that they found was that the time T required as a function of n (number of additions) could be expressed at any age level by the equation

$$T = Kn^m$$

where K and m are constants to be derived from the data and T and n have been defined. K and m were found to be related to age. K increased from about .88 in ages below fifty to 3.3 in ages above seventy, while m decreased from 1.26 below age thirty to about .98 above age seventy. When m equals 1.0, the time is in direct proportion to the number of additions per item. With m greater than 1.0, there is positive acceleration. A decrease in m , which changes very little with age, is more than offset by the increase in K , which increases with positive acceleration. The net effect is that the elderly take more time per addition as age increases.

The equation is given to demonstrate how searchingly performance on any test may be studied experimentally. In addition to this equation, the investigators derived an empirical equation for p (proportion of correct answers) as a function of n , and they found that parameters of this equation also changed systematically with age. A rational equation relating T , p , and n was developed. Analyzing the addition test logically into the components of perception of digits, addition of digits, carrying digits, and writing answers, they concluded that the elderly slow down in every respect but that decline in adding with age is largely a matter of weakness in organizing input information and relating it in relevant ways. This seems to describe cognitive defect rather than memory defect.

In general, paired-associates memorizing should be a matter of memory for implications, the second member of a pair being implied by the first member. Hulicka and Weiss (1965) have reported an experiment on the memorizing of figures paired with boys' names, with one group of subjects with mean age of thirty-eight and another with mean age of sixty-eight. The superiority of the younger group depended upon what kind of score was used, revealing some interesting age differences.

In runs of 15 trials for either group of subjects, the younger group made more correct responses, with scores of 92 versus 54. With equal exposure to material, then, the older S s were decidedly poorer. In a recall test twenty minutes later, the means were approximately 8 versus 4, in favor of the younger group. When the two groups were compared on the basis of trials-to-criterion scores, the younger group also did better, with means of 14 versus 21. But recall scores after twenty minutes were the same (7.6 and 7.6). This could be attributed to the fact that the elderly had had more exposure to the items, 50 percent more, in fact.

Further comparisons were made by matching young and old who had had the same mastery scores and presumably equivalent exposures to the input information. After a twenty-minute delay the means were about the same (7.8 and 7.6). After one week another test showed the older group remembering more, with mean scores of 6.5 and 7.5. The general picture is that the elderly were handicapped in acquisition of new information but that once they had acquired it, their retention was as good as that of the younger group. A correlation of .92 between learning scores and retention scores indicates the high degree of dependence of the latter on the former.

Divergent-production abilities In the preceding chapter much attention was given to the ages at which tested creative potential appeared to reach its maximum, and it was found that just past thirty seemed to be typical, in tests for ideational fluency and originality (Trembly, 1964a) and from the facts concerning the time for best creative output of distinguished people. With respect to decline, we have more analytical information.

Word fluency Tests calling for the rapid listing of words that satisfy some letter requirement have been common successful measures of factor DSU. With the King test, Bilash and Zubek (1960) found that means tend to keep at a high level to the age of

forty-five, then decline rapidly. With the Thurstone PMA test of Word Fluency, Schaie (1958) more or less confirmed this picture, with a curve showing a very slight rise to about age forty-three, then a decline at an increasing rate.

Birren (1955) has made a more searching study of the way in which the elderly react to this kind of test in comparison with young subjects. In each test, the examinees were told to write in two minutes as many words as they could beginning with a specified letter or letters. The initial letters were chosen so as to vary systematically the potential number of responses. The initial letters, with estimates of potential supply of such words, were *S*, 29,002; *C*, 22,909; *N*, 4,088; *Q*, 1,416; and *Gi* (no number given). An additional form of the test prescribed no initial letter or other restriction. There is an implied assumption that the potential supplies are the same for the elderly as for the young. This would depend very much upon each person's size of vocabulary. Since vocabulary declines little or none in relation to age, such an assumption is defensible.

One result was that the older subjects, whose ages ranged up to eighty-nine, gave fewer words in all test forms, regardless of potential supply. The elderly were relatively slower in the "easier" tasks, i.e., those with greater potential.

One interpretation of this result could be that with greater supply there is more interference in recall, as D. M. Johnson et al. (1951) have pointed out in connection with rapid successive recalls of items of a class of information from memory storage. The greater susceptibility of the elderly to interferences has been noted before, and it has been interpreted as confusion of information. The interference hypothesis is further supported by Birren's findings that the elderly gave fewer words relative to their writing speeds, which looks as if something related to recall were holding them back. They also gave more inappropriate responses, responses not fitting the prescribed class. This could also be attributed to confusion that cuts across classes, or it can be attributed to poor maintenance of set or to poor evaluation. Letting the inappropriate responses stand would suggest some role for evaluation.

Flexibility with respect to classes and transformations Although the last-mentioned idea would seem to suggest that the elderly err by slipping into wrong classes, there are other indications that they have difficulty in going from one class to another. The circumstances are somewhat different, however. The context here is in terms of classification tests, in which the person must shift from class to class to make a good score; the shifts of class are not slips or errors.

Thaler (1956) found that in the Weigel test elderly subjects tended to sort the objects in fewer classes and that some of the *Ss* were inclined to stick to the first classes they formed. This test should pertain to factor DFC, the divergent production of figural classes, or spontaneous flexibility with respect to figural classes. A result reported by Korchin and Basowitz (1956) may have similar significance. In a test in which successive pictures are shown, beginning with a cat, for example, and shifting gradually over into a dog, it takes the elderly longer to notice the shift. The change in cognition implies a shift in classification of the pictured object. The elderly also have relatively more trouble in seeing a greater change, as from a mouse to a car.

Some of the evidence we have with respect to decline curves for shifting of classes comes from Chown (1961). Five different tests for factor DMC, such as Brick Uses and Unusual Uses, were given to subjects of different ages. The five age curves were similar (thus giving replication with respect to rise and decline in the factor), rising slowly in the interval twenty to forty, remaining level from forty to fifty-five, and declining thereafter.

Other evidence comes from Bromley (1964), who gave tasks in the form of card sorting, and the Vygotsky and Shaw tests, in which *S* is to make successive classifications. Age brought systematic decrements in sorting output. The elderly showed relatively greater decrement in producing the more unusual classes, which suggests weakness in transformation ability, which should mean factor DFT. In fact, Bromley attributed the lower performance of the elderly to lack of transformation ability. But probably flexibility with respect to both classes and transformations was involved. Chown (1961) has reported a consistent rate of decline for a Match Problems test, a DFT test, with some acceleration.

Semantic fluency and originality From Trembly (1964a) we have information regarding a performance curve on a test like Consequences, which was described in Chapter 6. The examinees write an essay in response to the kind of question that begins: "What would happen if . . . ?" The score used by Trembly is the number of words written in the given time. It had been determined that a count of the number of words correlates very high with qualitative scores that require judgment. Summing obvious and remote consequences should give a measure of both DMU and DMT (ideational fluency and originality), with DMU dominant, because it is the writer's experience that the variance in scores for obvious consequences is about double that in scores for remote consequences.

Trembly's decline curves for both sexes show a maximum performance level being reached at about the age of thirty and a moderate rate of decline coming in by the age of thirty-five. By the age of sixty the mean score is equivalent to that obtained at about the age of eighteen, which is not nearly so great a decline as is found in many other abilities. As pointed out in the preceding chapter, the maximal score near the thirtieth year agrees very well with the Lehman findings that creatively productive people are most likely to create their most valued products in the early thirties or shortly thereafter, this event depending upon the accumulation of information, sometimes of a specialized kind.

Convergent-production abilities Convergent-production abilities, being less well known, have been much less investigated in connection with age. Garfield and Blek (1952) happened to give what is essentially a measure of factor NSR, the convergent production of symbolic relations, when they administered the Abstraction test of the Shipley-Hartford scale to teachers and teachers in training, in age groups in their twenties, forties, and sixties. The corresponding means were 31.6, 28.3, and 29.2. Only one of these differences was significant, and the declines were obviously small.

One of K. F. Riegel's tests (1959) given to old and young groups was a multiple-choice Antonyms test, and another was a multiple-choice Analogies test. Both should be expected to measure factors CMR and NMR, the latter if the examinee does much producing of related information before inspecting the alternative answers. The young group (mean age about nineteen) was significantly superior to the older group (above sixty-five) in both these tests.

There is indication of decline in tests of two other factors in the convergent-production category. Chown (1961) reported a very gradual decline from age twenty to age fifty in Hidden Figures, a measure of factor NFT (convergent production of figural transformations), then a rapidly increasing decline. Here is indication of decline in one kind of flexibility, of which there are a number of varieties. The other test that can be mentioned is Picture Arrangement, a measure of factor NMS (convergent production of systems), which Wechsler lists among his three tests showing most rapid declines.

Evaluation abilities There is some evidence of decline in a test called Perceptual Ability, a combination of King's tests that should measure factors EFU and ESU, both pertaining to evaluation of units, one figural and the other symbolic (Bilash & Zubek, 1960). It was reported that there were a slow decline to age forty-five and a more rapid one thereafter. Because the two factors are involved in the same test, it is difficult to decide how each factor's age curve would look if its test were scored alone. Since the two factors are parallel and contiguous in the SI model, the two curves might well be similar.

Friend and Zubek (1958) produced an age curve for the Watson-Glaser Critical Thinking Appraisal test. Although composed of several parts with somewhat different kinds of items, this test is most like a syllogism test, which is strongest for factor EMI (evaluation of semantic implications), with a secondary loading for factor EMR, a parallel relations ability. The curve was at its maximum in the twenties and into the thirties, after which there was a gradual decline. Even at seventy and older, 9 percent of the examinees scored above the mean at maximal ages. Estimation of part scores showed that the elderly tended to give more extreme T and F responses and fewer moderate ones (probably true and probably false), a result that was taken to mean some loss of flexibility. The elderly also tended to make relatively lower scores on items with personal reference as compared with those on impersonal items, which was taken to mean some loss in objectivity.

Welford (1964) made some comments about other apparent weaknesses in the area of evaluation. Older subjects seem to take more liberties in changing premises and in expanding upon their opinions. They sometimes have difficulty in selecting relevant items of information. And they show some loss of sensitivity to incompatibility of statements as well as permitting personal views to help determine conclusions.

In a different kind of test, but of a kind that has shown some loading on factor EMI and also on EMR (Nihira et al., 1964), K. F. Riegel (1959) has found that a younger group did better than an elderly one. A sample item reads:

A GRANARY always has: grain, elevator, cellar, mice, entrance.

The decision is to be made among alternative implications.

Decline in flexibility It is often observed that the elderly have lost in flexibility in various ways. Flexibility is not a single trait; so we should not expect individuals to be equally flexible or inflexible in every respect. Two kinds of flexibility in which declines have been demonstrated, flexibility in shifting from one class to another and flexibility in the sense of being ready to produce transformations, were mentioned earlier. Declines in these respects have been found in the production of classes and transformations. There are other indications of lowered flexibility.

Birren (1955) has reported that the elderly find it relatively more difficult to learn wherever there are conflicting associations and that learning to eliminate errors is likely to be more difficult than learning a new series of items. Entwisle (1959) has reported that in learning to drive a car the elderly are handicapped at first by conflicting earlier habits. Talland (1959) found that the elderly are disturbed by tasks that call for changes of set. As with other instances of weakness in coping with interferences mentioned earlier, these examples, also, may be confusions of information, which means that information is not sharply discriminated. Other instances may indicate a lack of readiness to produce transformations that are demanded by situations.

Complexity and systems Welford (1958) has remarked that the elderly show no notable decline in ability to have insights but that they show defect in maintaining condi-

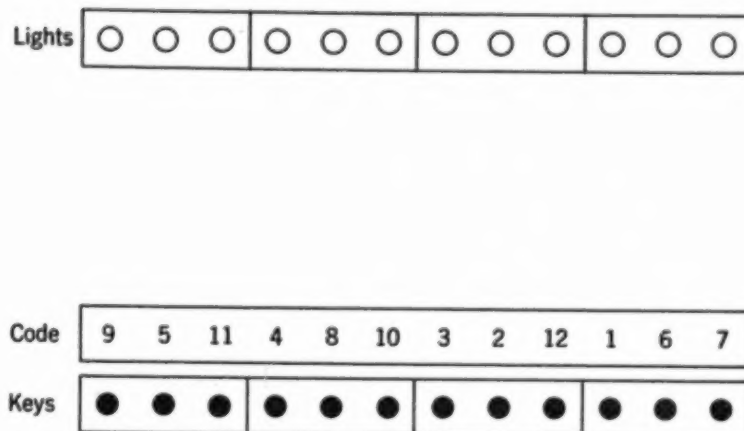


Fig. 18.3 Schematic diagram of arrangement of 12 stimuli and 12 response keys in a choice reaction-time experiment and the position of the code relating the two. (From Kay, 1954. *Reproduced by permission.*)

tions favorable for insight. Some insights depend upon keeping under consideration several items of information. Deficiency in this respect suggests low capacity for handling many items of information, whether this is a matter of short-term memory or is to be described by some other concept, such as the product of system. Other descriptions imply the latter interpretation.

Jerome (1962) observed that older individuals fail to understand methods that are described to them for dealing with tasks or reject methods that are demonstrated to them. Rejection under such circumstances may be due to lack of understanding, the subject having too much pride to admit that he does not understand. Methods are systems. Strategies are systems. Jerome also reports that the elderly lack clear analysis of the goal in problem solving, and they evidently lack good, clear search models, for their search behavior is often disorderly.

There are other indications that the elderly have much trouble because they cannot organize items of information into coherent wholes or systems. This phenomenon is also shown in reaction-time tasks. In a simple reaction-time task there is relatively little change in score in the ages between twenty and sixty (Welford, 1964), but there is pronounced decline when complexity is involved, as in discrimination reactions or choice reactions, when the number of alternatives increases.

A good example of reactions to a choice reaction-time task is found in Kay's experiment (1954). Twelve stimulus lights in a horizontal row (see Figure 18.3) were grouped by threes so as to be easily counted, number 1 being at the subject's left. The subject had a row of 12 response keys, but they were not related in direct correspondence with the stimulus lights. The stimulus lights were coded so that stimulus light 9 went with response key 1, stimulus light 5 with response key 2, and so on (see Figure 18.3). In the easiest arrangement, the code was provided for S's use directly above his response keys. In the next hardest arrangement the code was set midway between the lights and the keys, and in the hardest arrangement the code was laid next to the row of lights. The distance between lights and keys was 3 feet.

The results for the elderly and the younger groups were very systematic. Errors increased as an accelerated function of both age and difficulty, as represented in Figure 18.4. The main source of difficulty was the complexity of the system of light-key relationships.

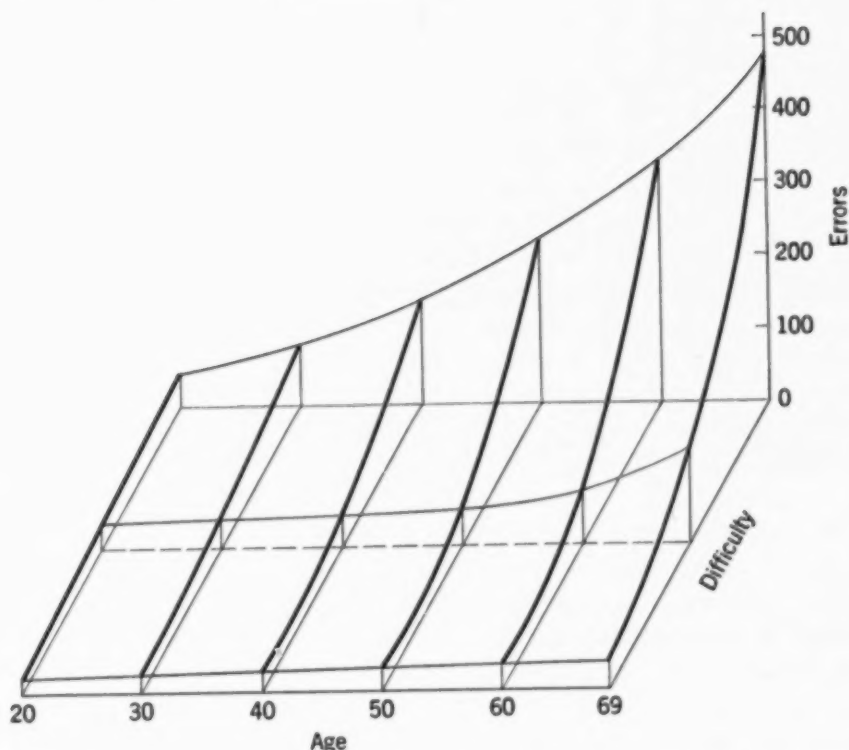


Fig. 18.4 Errors in a choice reaction-time experiment as functions of age and difficulty level. (From Kay, 1954. Reproduced by permission.)

The oldest group made four times as many errors and took five times as long as the subjects below the age of twenty-five in the most difficult task. The age differential decreased considerably for the simpler tasks. If we assume that sensory time and motor time remained fairly constant under all arrangements, the increasing time must be attributed to central brain processes apart from the primary sensory and motor centers.

Clay (1954) has demonstrated the effects of complexity with age by means of another type of task. Given counters with digits 1 to 4 on them, *S* was to enter them in the cells of a small matrix so as to achieve given sums for the rows. The task was made more complex by varying the order of the matrix from 3×3 to 6×6 . Elderly subjects showed relatively more errors for the more complex tasks. They could often see that they had made errors but did not seem to know what to do about them. This observation suggests some strength in evaluation but weakness in divergent production.

In another kind of task, involving problem solving of an inductive type, *S* was to discover which keys turned on which lights (Clay, 1957). From age twenty to age eighty, time was an increasing function of age. After reaching a maximum, however, the average time again dropped off as some of the subjects gave up the task. *Ss* gave up when they became confused. The number of errors rose more sharply after the maximum time had been reached. The more complex the task, the earlier the age of giving up. Clay's hypothesis was that *S* takes more time in order to handle the additional information with which he must deal. But added time means more dependence upon short-term memory, which has definite limits. Another interpretation is that the more complex problem requires the development of a more elaborate system, which is beyond the individual's system-construction powers. Systematizing input, in any case, reduces the number of items

of information with which the individual has to deal and thus keeps the input within his capacity for handling it.

Some general conditions of decline

Considerably less attention has been given to the determiners of decline of intellectual abilities than to their growth, but as populations of the elderly increase, the problem of how to prolong mental fitness and to prevent senescence becomes more important. A knowledge of determiners is of first importance in this connection. Many of the same conditions that promote or retard intellectual development are naturally suspected of having some bearing on the maintenance of intellectual status and perhaps on retarding its decline. As in the last chapter, these conditions will be examined under the headings of heredity and the brain and of various environmental determiners.

Heredity and the brain Not a great deal is known about the relation between heredity and mental decline. But since heredity helps to determine development, it also determines the status from which decline starts. The new problems would be concerned with rates of decline and with the possibility of senility before death.

Falek, Kallman, Lorge, and Karvik (1960) found that elderly one-egg twins show greater similarity of longevity than do two-egg twins. On the one hand, this might suggest that the same hereditary disposition that leads to longevity also provides for prolongation of better status in mental powers. On the other hand, this might have the indirect effect of providing opportunity for senility if the brain gave out before other organs. There is a little indication in favor of the first of these two contingencies. Jarvik et al. (1962) found that the survivors of pairs of twins had higher mean scores when tested past the age of sixty than did their mates who died. In both the studies mentioned in this paragraph, it was found that one-egg twins make more similar scores in tests than do two-egg twins. All that can be said is that at least some of the similarities of one-egg twins found during development appear to persist in old age.

Some of the brain conditions that affect development of intellectual status are known, as mentioned in Chapters 15 and 17. We need to consider whether any of them play any special roles during decline. One of the variables is the amount of brain tissue available. Many years ago, Pearl (1905) found a curve for decline in human-brain weight that looks very much like that for general intellectual decline, as in Figure 18.1. But other physical decline curves could probably be found to parallel either of them. If loss in brain weight is highly significant, how does it bring about loss of intellectual ability? In this connection, we are reminded of Lashley's classical finding that in rat-maze learning it does not matter where the cortical loss occurs but that it does matter how great the loss is.

The loss of human brain cells by death has been noted before. How does loss of brain cells affect functioning? Welford (1958) has theorized that loss of brain cells lowers "informational capacity." This is supposed to reduce the strength of communication signals, increasing the "random" brain activity or "noise." Or, the absolute amount of noise may remain constant, but with loss of cells the average noise per cell is greater. In either case, there is a decrease in signal-to-noise ratio, with danger of overloading circuits, causing momentary lapses of attention.

Such a theory is clearly derived from concepts in communication engineering, which is not a bad source of analogies where brain functioning is concerned. The reduced capacity for input of the elderly has seemed to be descriptive of a number of their weaknesses. There may be a reduced capacity for productive activities as well. But there seems to be

another feature that is just as important, and that is the reduction in clarity of discriminations, which opens the way for interferences and confusions. Whatever condition of the brain that permits clear discriminations should be found to decline in natural aging.

After considering reduced speed of nerve conduction as a possible source of mental decline, Birren (1955) rejected that condition as being very important. He has considered the number of cells available for use in any operation and conditions of nerve fibers as possible determiners (Birren et al., 1954). Lowered oxygen supply is often mentioned as a suspected determiner of aging, for with anoxia brain cells die or their chemical composition changes (Bondareff, 1959). It has been observed that pigment accumulates in the brain cells, but just how this condition affects their functioning is not known. Possibly it affects the cells' replicating activity in the formation of proteins that are needed in mental functioning. Such a condition would make difficult the formation of phase sequences in Hebb's sense.

Environmental and other conditions

Decline at different levels of ability Much interest has been expressed in the rates of decline and the ages of beginning of decline for individuals of different intellectual status. The variable of intellectual status should mean standard assessment of all individuals at some particular age, for example, at the time of young adulthood. A really satisfactory study of this problem would require a longitudinal investigation of a large sample of wide range of status over the years of adulthood to old age. No study of this kind has come near these specifications or is in any way comparable to the Berkeley growth study (Bayley, 1949). We shall have to be content at present with some piecemeal attempts, with groups followed over limited numbers of years.

There has been a more or less common hypothesis that the bright young people stay bright longer and the dull young people start to decline earlier. This principle should show up in the form of increasing dispersions of scores with advancing age. There are two cross-sectional studies that bear upon this hypothesis. H. E. Jones (1959) cited a study by Raven with his Progressive Matrices test, in which he compared scores at centiles 2, 50, 75, and 95 at ages five to sixty. The trend curves as functions of age did diverge, but more during the years before twenty than after that age. This finding was confirmed by Foulds (1949) with the same test and also with the Mill Hill Vocabulary Scale in a range of sixteen to sixty-five.

In Terman's gifted group, Bayley and Oden (1955) applied the Terman Concept Mastery Test at the average age of 41.5 to some of the individuals and to their spouses, all of whom had taken the same test at the average age of 29.5; in other words, retested them after an interval of twelve years. There was an average increase of about $\frac{1}{2}$ standard deviation for both men and women and for all levels of ability within the sample, which had been selected during childhood with a minimum IQ of 140. There were some slight indications of a regression effect for the very highest individuals, but it was suspected that the test did not provide enough scope for these cases. There was no group with initially average IQ or with initially low IQ with which to compare the gifted group. The result for the gifted group might have been nothing more than the usual gain in retesting, particularly with a vocabulary type of test in which there is often gain as a function of age.

According to the general hypothesis, mentally deficient groups should show more than average decline in tests. Clare W. Thompson (1951) also tested the hypothesis that declines should begin earlier in the mentally deficient than in the intellectually normal. She reasoned that since the mentally defective tend to achieve their maximum scores at earlier

ages and since their life-spans are shorter than those of normal individuals, they live a kind of telescoped existence. They should start a decline earlier in intellectual abilities.

Thompson compared a group of morons whose IQs were in the range 50 to 70 on the Stanford-Binet scale with a group of normals. In the interval of ages twenty to thirty, the mentally defective groups showed declines in all 10 of the performance tests (from the Wechsler scale and other sources), whereas the normals showed declines in only 4, in a cross-sectional study. The hypothesis of the earlier declines for the mentally deficient thus appeared to be supported. In two verbal tests, however, the mentally deficient did not show declines in their twenties; so the hypothesis was not supported for all kinds of abilities.

A study that gave results counter to the expected differential decline of the mentally deficient was reported by Bell and Zubek (1960). Four groups of mentally deficient subjects at mean ages of twenty, thirty, forty, and fifty were retested with the Wechsler scale after five years. The changes in IQ at the successive age levels were the following gains, respectively:

Total-scale IQ	9.1, 7.6, 5.7, 4.4
Verbal IQ	4.1, 4.2, 4.3, 4.4
Performance IQ	12.0, 9.3, 7.4, 1.7

It was pointed out that these trends are similar to those found in other longitudinal studies with normal groups, in contrast to the results of cross-sectional studies. The fact that the verbal tests held up better in the older groups is also consistent with other findings.

Possible reasons for the gains were considered. Practice effects were discounted in view of the low IQs of the subjects. The same administrator gave the tests in both instances. There was an improved environment that applied to some cases, in the form of moving to a better institution. It was admitted that there could have been some underestimates of IQ on the first testing. Even considering all these incidental possible determiners, the loss in IQs for the mentally deficient population expected from the hypothesis did not seem to apply. We have a long way yet to go to determine changes in intellectual level in all respects as functions of characteristic status.

Education There is no doubt of the strong correlation between amount of formal education and *development* of intellectual abilities. Is there also a bearing of amount of education upon the rate of decline of abilities? To the extent that individuals maintain their rank orders from maturity to old age in tests of intellectual abilities, of course, we should expect the correlation between education and status in old age to remain strong. The relation to *rate* of decline is a different problem, and it is akin to the problem of the relation of rate of decline to mature status in the abilities, because of the confounding of education and intellectual status.

Education is becoming more and more hazy as a variable, as the institution of adult education and opportunities for informal education become more general. Definitive studies with the education variable must pay much attention to the content of the education and to its quality as well as its quantity. To the extent that individuals keep up with educational activities they should be exercising many of the intellectual functions, and exercise should contribute to keeping abilities at higher levels.

That highly educated persons are not exempt from some degree of decline is suggested by a study by Sward (1945). His subjects were university teachers, active and retired. One group were of ages twenty-five to thirty-five with a mean of thirty-one, and the others were of ages sixty to eighty with a mean of sixty-six. There was a matching on other

variables. In seven of eight tests used, the younger group was significantly superior. The seven tests were Ingenuity, Problems, Symbol Digit, Number Series, Word Meanings, Analogies, Arithmetic Problems, and Artificial Language. On the Synonym-Antonym test the older group was significantly superior. It is unlikely that the difference on the last-mentioned test was due to CMU variance, since the difference was in the opposite direction to that on the Word Meanings test, which is more likely to measure CMU. On that test, however, the difference was significant only at the .05 level. Within the two groups, with their relatively short ranges of ages, the correlations of test scores with age were generally small but tended to be stronger (negatively) for the older group, indicating more rapid decline in that group. It was concluded that the elderly were more impaired with respect to rate of work than with respect to accuracy or quality.

Occupation Individuals who engage in occupations that exercise intellectual functions should keep themselves at higher levels, depending upon the demands for those functions in their work and in their occupational milieu. Foulds (1949) had an opportunity to test this hypothesis in a limited way in his investigation with the Raven Progressive Matrices test and the Mill Hill Vocabulary Scale with two occupational groups, engineers and manufacturing employees who were concerned with photographic equipment. Subjects were in groups with ages sixteen to sixty-five. Foulds compared decline curves for the two kinds of personnel in the two tests, one of figural content and one of semantic content, the latter being more specifically for factor CMU. The Raven score is probably weighted with factors CFR and NFR.

For the engineers, the average Raven score was maximal between the ages of fifteen and twenty; for the factory workers, who presumably had inspectional and manipulative types of assignments, the maximal mean score came between twenty and twenty-five. After thirty the declines were similar for the two groups, but the factory workers kept a slight lead to the age of fifty-five. In the vocabulary test, engineers came to their highest average between forty and forty-five and the workers between thirty-five and forty. There were similar, slow declines after forty, with the engineers keeping a slight lead that narrowed to the age of sixty.

It is tempting to attribute the differentials in ages of maximal status and in retaining superiority during decline to the fact that engineers were of necessity keeping up development in acquiring new verbal concepts while the factory workers were practicing skills involving figural abilities. But it should be noted that declines set in for the two groups at about the same time in both tests and that their mean scores converged to about the same level at age sixty-five. It may be that both groups had developed as far as was needed rather early in their careers and that they had acquired specific skills that carried them through. The generalized skills could then lapse normally without undue handicap in their work.

That occupational experiences may have bearings upon keeping certain relevant aptitudes strong or even upon improving them is suggested by some limited data presented by Glanzer and Glaser (1959). Air Force Reserve pilots and commercial-airline pilots who had taken the Air Force classification battery during World War II were retested thirteen years later with four of the same tests. The later mean scores were significantly higher (at the .001 level) in all four tests. The tests, their known factor contents, and the first and second means were as follows:

Instrument Comprehension (CFS-V, spatial orientation): 34.8, 47.4
Mechanical Principles (CFT, visualization): 18.7, 25.4

Numerical Operations (numerical facility): 69.1, 83.1
 Spatial Orientation (EFU, perceptual speed): 29.6, 32.9

All except Numerical Operations are known to have relevance for prediction of learning to fly an airplane. It is possible that the pilots' later work involved some computational exercise. Some of the gains may be the general increments such as are found in retesting, even after a period of years, in longitudinal studies. Control groups would be needed to determine how much could be attributed to the pilots' occupational experiences.

Gains in measures of factors consequent to educational experiences, as cited in Chapters 12 and 14, should lead us to expect benefits from occupational experiences that provide exercises in tasks in which the factors are relevant. There is no available experimental information of this kind for middle-aged and elderly populations.

Summary

Investigation of decline of intellectual status with age presents more numerous difficulties than the study of development, owing to the multitude of variations in life circumstances that interfere with experimental control. Cross-sectional and longitudinal studies have been made, but the two approaches tell somewhat divergent stories. Whereas the former show almost universal declines in test performances among older groups, the latter more often than not show gains, although longitudinal studies have not been numerous, especially in the later years of life.

There is evidence that the factorial structure pertaining to certain factors remains about the same in groups of individuals who are aging normally, with only a slight indication that some simplification of structure occurs later in life, as shown by slightly higher intercorrelations of tests. Decline in the case of senile dementia is qualitatively different from normal decline. For the senile population, intercorrelations of tests rise markedly, suggesting changes in factor structure.

Most of the research has been in terms of changes in scores from special tests rather than in a global score, for differential rates of decline are numerous and striking. Not enough of the range of structure-of-intellect factors has been explored in relation to adult ages to give rise to many general principles related to SI categories. Within any one operation category, decline curves based upon test scores are quite varied. Scores that measure factor CMU, verbal comprehension, have represented an outstanding exception, by showing continued gains even into the sixties and by showing late-starting declines. Other semantic-cognition abilities, however, show more typical declines.

Two kinds of SI abilities stand out with respect to defects in the elderly. One has to do with the formation and use of systems of information, which is manifest in tests involving complex operations. An alternative hypothesis is that the older person's capacity for information intake and perhaps, also, for handling information in productive thinking is more limited. In the case of input, this limitation may depend upon a weakness in short-term memory.

The other notable area of weakness is in the abilities involving flexibility, flexibility with respect to shifting between classes and with respect to transformations in information. A rather general weakness is susceptibility to interferences and inability to correct errors. Interferences go with lack of clear discriminations and, therefore, with loss of information.

Little is yet known about the causes of intellectual decline in the various respects. Heredity appears to have a bearing, but possibly largely because it has a bearing upon

development. The most suspected brain condition in connection with decline is inadequate oxygen supply, as a consequence of which brain cells change chemically and may die. Exercise of brain cells by virtue of continued education, formal and informal, including that provided by occupational activity, may have halting effects, but this hypothesis needs much better experimental investigation.

ERIC Clearinghouse

SEP 14 1968

on Adult Education

Contents

<i>Preface</i>	vii
 I Introduction	
1 <i>Historical background</i>	2
Development of mental tests	2
Conceptions of intelligence	11
Intelligence as learning ability	14
Summary	20
2 <i>The investigation of intelligence</i>	21
Intelligence in experimental psychology	21
Genetic approach	22
Multivariate approach	25
Advantages of the factor approach	35
The psychological nature of intellectual factors	41
Summary	44
3 <i>General theory of intelligence</i>	46
Frames of reference and theory	46
Factor-analytic models and methods	48
Spearman's <i>g</i> and <i>s</i>	56
Hierarchical models of factors	57
The structure-of-intellect model	60
Other structural theories of intelligence	66
Summary	67
 II The structure of intelligence	
4 <i>Cognitive abilities</i>	70
Cognition of units	71
Cognition of classes	80
Cognition of relations	85
Cognition of systems	91
Cognition of transformations	100
Cognition of implications	104
Other possible cognition factors	108
Summary	108
5 <i>Memory abilities</i>	110
Early evidence for memory factors	110
Some special problems in memory testing	111
Memory for units	114
Memory for classes	119
Memory for relations	121
Memory for systems	125

Ac 002 495

Memory for transformations	128
Memory for implications	132
Summary	136
6 Divergent-production abilities	138
Divergent production of units	139
Divergent production of classes	143
Divergent production involving relations	146
Divergent production of systems	148
Divergent production of transformations	152
Divergent production of implications	159
Some special problems with divergent-production abilities	161
Summary	169
7 Convergent-production abilities	171
Convergent production of units	172
Convergent production of classes	173
Convergent production of relations	174
Convergent production of systems	175
Convergent production of transformations	178
Convergent production of implications	181
Additional comments	183
8 Evaluative abilities	185
Evaluation of units	186
Evaluation of classes	191
Evaluation of relations	193
Evaluation of systems	195
Evaluation of transformations	198
Evaluation of implications	200
Summary	202
9 Categories of operation	203
Cognition	203
Memory	211
Divergent production	213
Convergent production	214
Evaluation	217
Summary	220
10 Categories of information	221
The nature of psychological information	221
Categories of content	225
Products of information	238
Psychology and logic	245
Summary	249
 III Implications for psychological theory	
11 Perception and cognition	251
Input operations	251
Conscious and unconscious information	257
Perceptual development	261
Summary	267

12	<i>Learning</i>	268
	What is learned?	269
	How products are learned	271
	Roles of factors in learning	279
	The role of frequency in learning	289
	Reinforcement as information	292
	Executive functions	293
	Summary	294
13	<i>Retention and recall</i>	296
	Retention	296
	Recall	302
	Summary	311
14	<i>Problem solving and creative production</i>	312
	Similarities between problem solving and creative production	312
	A general model for problem solving	313
	The nature of problem solving and creative production	316
	General conditions affecting creative production	331
	Computer simulation of thinking	342
	Summary	344

IV Determiners of intelligence

15	<i>Physical basis of intelligence</i>	347
	Heredity	347
	General features of the human brain	356
	Some special features of the brain	359
	Intellectual functions and the brain	366
	Chemical conditions	381
	Summary	385
16	<i>Environmental and other conditions</i>	387
	Socioeconomic status	387
	Environmental deprivations	392
	Environmental enrichments	396
	Education and intellectual development	400
	Some other conditions	403
	Summary	410
17	<i>Intellectual development</i>	412
	Psychometric descriptions	412
	Qualitative descriptions	426
	Summary	437
18	<i>Intellectual decline</i>	439
	Research methods	439
	Decline in composite scores	445
	Decline in intellectual factors	447
	Some general conditions of decline	457
	Summary	461

V Reflections

19	<i>Retrospect and prospect</i>	464
	Preceding content	464
	Factor-analytic investigation of intelligence	469
	The future of intelligence testing	471
	The future of education	473
	<i>References</i>	478
	<i>Name Index</i>	511
	<i>Subject Index</i>	523